

生産関数フロンティア分析と統計的因果探索を組み合わせた道路インフラ整備効果の解明

国土交通省近畿地方整備局 正会員 ○安原優樹
 広島大学 正会員 塚井誠人

1. はじめに

社会基盤整備事業効果を長期予測する手法は、予測者が分析を行う年次と予測対象年との間に生じる技術革新等の事象や他社会基盤整備事業の波及効果等の影響を受ける。その結果、過去に行った事業効果の現在の予測値と、現在の観測値について、常に乖離が生じるという課題がある。他方で国土交通省は、発現した様々なストック効果を把握し、可視化することを推奨している¹⁾。そのためには、事業効果の予測モデルと現実の観測値との乖離が生じる原因をモデル構造に遡って検討し、新たな経済モデルを開発する手順が必要である。本研究では、生産関数フロンティア分析を用いて地域の下限総生産を算出し、観測総生産との残差に統計的因果探索を適用することで、生産関数で説明できない付加価値への寄与要因を明らかにする。これにより、経済モデルが想定した因果構造の妥当性を検証する。

2. 統計的因果探索

統計的因果探索とはデータから因果グラフを推測する手法である。本研究では、統計的因果探索の中でも関数形に線形性、誤差変数に非ガウス分布、因果グラフに閉路を持たない非巡回仮定を置くLiNGAMモデル(liner non-Gaussian acyclic model)²⁾を用いる。p個の観測変数に対するLiNGAMモデルは式(1)で示される。

$$x_i = \sum_{j \neq i} b_{ij} x_j + e_i \quad (i = 1, \dots, p) \quad (1)$$

式(1)の変数はx: 観測変数, b: 係数, e: 誤差変数である。式(1)を行列式で表すと式(2)のように書ける。

$$\mathbf{x} = \mathbf{B}\mathbf{x} + \mathbf{e} \quad (2)$$

式(2)の係数行列Bのゼロ・非ゼロパターンから因果の順序や強さを推計する。

表-1 正しい因果的順序が推計された割合

サンプル数	推計割合[%]
10000	100
1000	97.7
100	40.1

表-2 正しい因果的順序が推計された割合

係数行列の要素 b_{ij} の条件	推計割合[%]
$0.05 \leq b_{ij} \leq 0.10$	62.4
$0.10 \leq b_{ij} \leq 0.20$	99.6
$0.20 \leq b_{ij} \leq 0.50$	100
$0.50 \leq b_{ij} \leq 1.00$	99.9
$1.00 \leq b_{ij} \leq 1.50$	99.4
$1.50 \leq b_{ij} \leq 2.00$	96.4

3. 正しい因果的順序が推計される条件

インフラストック効果の推計では、その効果が空間的にスピルオーバーする³⁾ため、実効サンプル数は見かけのサンプル数よりも小さくなり、パラメータ推定が不安定になるおそれがある。そこで、サンプル数と因果構造探索の関係について、数値シミュレーションを行った。まずLiNGAMの仮定をすべて満たし、因果的順序が明確な6変数の擬似データを、乱数を用いて作成する。そして、作成したデータサンプルから因果的順序を導出する試行を1000回繰り返し、正しい順序が得られた割合を算出する。サンプル数ごとの結果を表-1に示す。

同表より、正しい因果的順序が推計されるには1000サンプル以上が望ましいことが分かる。また正しい因果的順序が推計される条件として、係数行列Bの要素 b_{ij} のスケールが考えられる。この理由として、次のようなトレードオフが想定できる。独立成分分析(Independent Component Analysis, 以下ICA)²⁾を用いたLiNGAMアプローチでは、回帰成分のスケールが独立成分のスケールに対して大きいと、有効な独立成分の検出が難しくなる。逆に小さいと、独立成分の僅かな攪乱の影響を受けて混合係数の推定が

キーワード: LiNGAM, ストック効果, 事後的検証

連絡先: 東広島市鏡山1-4-1 広島大学大学院先進理工系科学研究科 塚井誠人

yasuhara-y86eb@mlit.go.jp, mtukai@hiroshima-u.ac.jp

不安定になる。つまり、どちらにせよ ICA から正しい因果的順序を得ることが難しくなると予想される。 b_{ij} のスケールの条件について得たシミュレーションの結果を表-2に示す。結果は想定を満たしており、因果的順序が正しく推計できる範囲が存在していた。この係数のスケールの範囲を基に、因果構造を推計した際に算出される決定係数の妥当性を評価する。

4. データの概要

(1) 下限フロンティアと観測総生産の残差

地域の下限生産額を算出する生産関数フロンティアとして、式(3)の生産関数⁴⁾を用いる。

$$\ln(y_n) = f(K_n, L_n; \beta) + v_n + u_n \quad (3)$$

ここで、 y_n ：地域 n の域内総生産額、 K_n ：地域 n の投入資本額、 L_n ：地域 n の労働量、 β ：パラメータ、 v_n ：攪乱項、 u_n ：効率性を表現する確率項である。この u_n が下限フロンティアと観測総生産との残差であり、以降「eff_cd」と表記する。

(2) 使用データの詳細

本研究で使用したデータを表-3に示す。アクセス指標の定義は式(4)に示す。

$$AC_i = \sum_{j=1}^n \frac{P_j}{t_{ij}^2} \quad (4)$$

ここで、 i ：対象地域（市町村）、 j ：市町村役場あるいは都道府県庁、 t_{ij} ： i から j への自動車所要時間、 n ：地域数、 P_j ：地域 j の人口である。

また、本研究では産業別の生産額への寄与を明らかにするため、従業者数と事業所数を基盤産業と非基盤産業に分割した。その分割方法は、1)中分類した各産業の従業者数を地域の人口で除する、2)その値を産業ごとで標準化する、3)全市町村×全産業の中で上位3分の1を基盤産業、残りを非基盤産業とする、4)従業者数の基盤・非基盤を基に事業所の基盤・非基盤を割り当てる、という手順で求めた。

5. 因果グラフとその考察

表-3のアクセス指標について、終点を市町村役場と都道府県庁の2通りの場合でそれぞれLiNGAMを適用して結果を比較したところ、後者のほうが、推計モデルの精度ならびに符号条件とも、良好だった。後

表-3 使用データの詳細

年次*	平成22年, 平成27年
範囲	国内の約1700市町村
変数***	人口/国勢調査 事業所数/経済センサス 従業者数/経済センサス アクセス指標/NITAS eff_cd (下限フロンティアとの残差)

*観測年次が異なるデータは直近の平成21年, 平成23年, 平成26年, 平成28年の値を用いる
**以下はデータの出典を表す

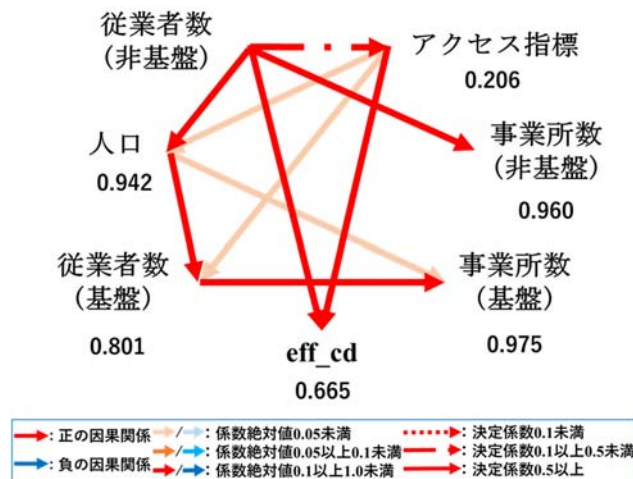


図-1 推計された因果グラフ

者を基に推計した因果的順序に基づいて作成した因果グラフを図-1に示す。なお図-1のモデルは、表-2の係数条件を満たしている。同図より、下限フロンティアとの残差を表す「eff_cd」には、非基盤産業の従業者数とアクセス指標が正の影響を与えている。

6. 結論

本研究では、LiNGAMに基づいて、社会基盤整備のストック効果の分析手法に関する事後的検証手法を提案した。その結果、下限フロンティアとの残差には、非基盤産業の従業者数とアクセス指標が正の因果関係が確認された。また、因果的順序の妥当性の評価条件を新たに発見した。本研究の課題は、産業の基盤・非基盤の分割基準やより妥当なアクセス指標の検討、小標本時の推計結果の安定性確保である。

参考文献

- 国土交通省・社会資本整備審議会計画部会専門小委員会：ストック効果の最大化に向けて～その具体戦略の提言～,2016
- 清水昌平：機械学習プロフェッショナルシリーズ統計的因果探索, 講談社(2017)
- 塚井誠人, 江尻良, 奥村誠, 小林潔司：社会資本の生産性とスピルオーバー効果, 土木学会論文集 No.714/IV-57, pp53-67, 2002.
- Subal C. Kumbhakar, Ca Knox Lovell: Stochastic Frontier Analysis: An Econometric Approach, Cambridge University Press, 2000